INEEL/CON-99-00199
PREPRINT

The VirtualwindoW: A Configurable, Modular, Stereo Vision System

R. A. Kinoshita
M. O. Anderson
M. D. McKay
W. D. Willis

SPIE 13th Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls

The VirtualwindoW: A Reconfigurable, Modular, Stereo Vision System

Robert A. Kinoshita, Matthew O. Anderson, Mark D. McKay, Walter D. Willis

Lockheed Martin Idaho Technologies Company
Idaho National Engineering and Environmental Laboratory Robotics Group
P.O Box 1625, Idaho Falls, ID 83415-2220

ABSTRACT

An important need while using unmanned vehicles is the ability for the remote operator or observer to easily and accurately perceive the operating environment. A classic problem in providing a complete representation of the remote work area is sensory overload or excessive complexity in the human-machine interface. In addition, remote operations often benefit from depth perception capability while viewing or manipulating objects. Thus, there is an on going effort within the remote and teleoperated robotic field to develop better human-machine interfaces. The Department of Energy's Idaho National Engineering and Environmental Laboratory (INEEL) has been researching methods to simplify the human-machine interface using atypical operator techniques. Initial telepresence research conducted at the INEEL developed and implemented a concept called the VirtualwindoW. This system minimized the complexity of remote stereo viewing controls and provided the operator the "feel" of viewing the environment, including depth perception, in a natural setting. The VirtualwindoW has shown that the human-machine interface can be simplified while increasing operator performance. This paper deals with the continuing research and development of the VirtualwindoW to provide a reconfigurable, modular system that easily utilizes commercially available off the shelf components. This adaptability is well suited to several aspects of unmanned vehicle applications, most notably environmental perception and vehicle control.

Keywords: virtual presence, stereo vision, system integration, reconfigurability, modular design, hands-free control, Commercial Off-The-Shelf (COTS)

1. INTRODUCTION

Current stereo vision technology provides the ability for a human to perceive in a clear and accurate manner a remote area that preserves depth perception. Since 1995, the robotics group of the INEEL has been developing a system called “VirtualwindoW” which now provides the capability to quickly deploy vision and actuation technologies, including stereo vision in a wide variety of applications. From its initial incarnation as a single-purpose research platform, the VirtualwindoW has grown into a general-purpose device that is easily extensible using commercial components.

2. PREVIOUS WORK

2.1. Initial Development

The VirtualwindoW was first developed during a research project to create and evaluate the use of telepresence technology on a remote vehicle platform. To do this, a stereo camera set was integrated with a Remotec Andros tracked robot chassis. To provide versatile viewing of the operating environment, the cameras were mounted on a high-speed pan-and-tilt (PTU) device. The PTU allowed the camera to view the entire area surrounding the robot. To perform general-purpose tasks, a multi-axis manipulator arm and gripper were also mounted on the Andros vehicle.

In order for a single operator to successfully operate this machine, two innovative control techniques were implemented that are still in use on the current VirtualwindoW platform. First, it was found that it would be useful for an operator to be able to move the camera field of view while simultaneously performing other operations (i.e. moving the robot or picking up an object with the gripper). This was implemented through the use of a head-tracking device. The position and orientation of the operator’s head could then be used to move the camera in an intuitive manner. For example, a head turn to the right would cause the camera to pan the view to the right. The overall affect is that the movement of the cameras are controlled by the operator’s gaze. There were also some miscellaneous operations that were provided through voice recognition technology. Switching video inputs to the screen and temporarily suspending robot motion are two examples of such functions. This initial work was presented to the International Society for Optical Engineering Symposium (SPIE) in 1996[1]. More details can be found in the proceedings of that meeting.

2.2 Dual-Arm Work Platform (DAWP) Field Deployment

An opportunity to continue development of the VirtualwindoW platform came with its incorporation into the Dual-Arm Work Platform. This device (See Figure 1) was constructed to aid in the dismantling of the CP-5 nuclear reactor at the Argonne National Laboratory-East. Two large robot arms are mounted on a metal framework controlled by an operator station located safely away from the reactor. VirtualwindoW provided a way for the operator (whose hands are busy controlling robot arms) to use and adjust the two stereo camera sets utilized in observing the work area. The depth-perception capability was also useful in accurately positioning tools to perform the demolition work. In preparation for this task, system hardware and software were both improved.

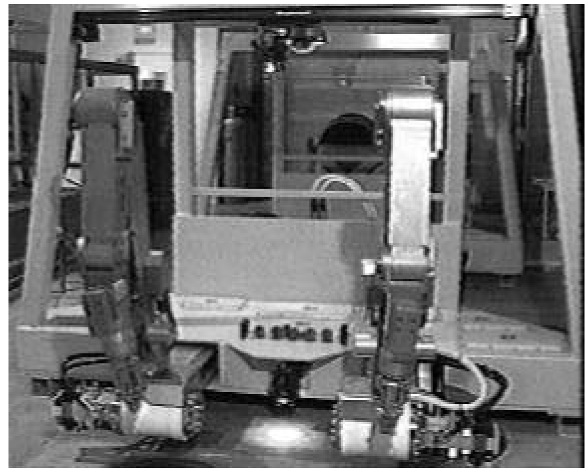

Figure 1 Dual Arm Work Platform

The VirtualwindoW hardware required changes in packaging, device support, and computing to be ready for integration. In its initial form, the system comprised three computers and associated equipment arranged on a table. The system was made compact and modular by moving all functions to one IBM PC-comptible computer, which was mounted in a single enclosure. This enclosure fits into a standard 19-inch rack, similar to that already in use for other DAWP control station components. Stereo vision was provided by a CrystalEyes system, which was also rack-mounted. The CrystalEyes system provides the capability to multiplex the two stereo video signals on a single video display. Special operator glasses are then used to convert these video signals into depth perception enabled video. There was also a need to add a second pan-and-tilt unit mounted on a long linear actuator. This “bird’s eye” placement allows the camera to be moved from side to side to look ‘around’ an obstacle. All hardware items (the new computer, PTU, and slider bar) were purchased off-the-shelf and integrated together. The resulting improved control station is shown in figure 2.

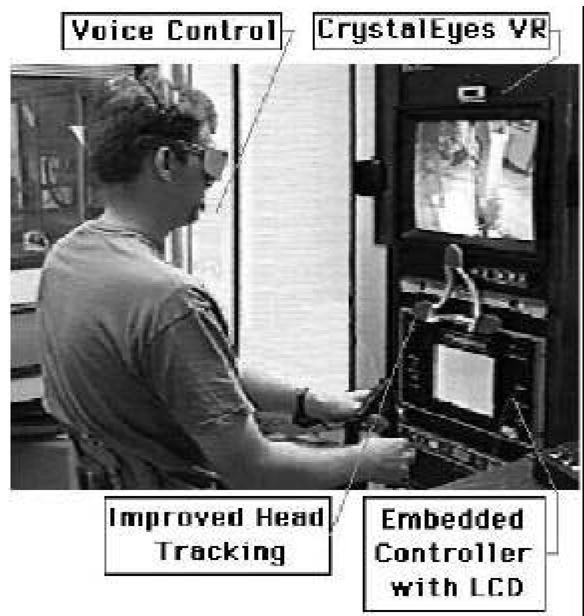

Figure 2 DAWP Control Station

In the course of merging all system software (including speech recognition and head-tracking functions) into one computer, several other improvements were made. First, the PTU control software was modularized so that multiple PTU devices could be simultaneously connected (previously there was only allowance for one). Support was also added to control the slider bar through the new head motion of shifting from side to side. Most significant, though, was the creation of the zone concept. Each head motion (turn left/right, tilt up/down, slide left/right) was divided into sections, including a dead zone and calibration capability. As an operator’s head inclines farther from the center dead zone, motion speed increases. This eliminated the requirement for precise operator head control to keep cameras steady. The DAWP deployment work was presented in detail to the American Nuclear Society in 1997[2].

3. LATEST GENERATION DEVELOPMENT WORK

3.1 Philosophy

The next opportunity to improve the VirtualwindoW system came in 1998. Through the Laboratory Directed Research and Development (LDRD) program of the U.S. Department of Energy, the system was modified to add a stereo vision zoom capability. In the course of completing the scope of the LDRD work, several other long-needed improvements were made to increase usability and flexibility. In order for the VirtualwindoW to be quickly applied to a wide variety of situations the system was internally modularized and made to be reconfigurable. New hardware support was also added to the system software, and the number of serial ports on the computer was increased to eight.

3.2 Flexibility

As the variety of hardware devices connected to the VirtualwindoW increased, it became progressively more complicated to fit them into the existing software architecture. Before now, any changes required changing the system software and rebuilding it, which was more appropriate to the earlier research phase of the program. The software was rewritten to allow the system to be configured through the use of a file that specifies which hardware is to be used (i.e. PTU, head-tracking device, etc.) and to what serial ports they are connected. All that is required to change the system behavior is to stop the program, edit the file, and restart it. The concept of camera switching was also expanded to a new concept called a “View”. Now up to five separate views can be specified, each one having its own video and device setup that can be switched by voice (if in use) or keyboard commands. Connected hardware devices can be associated with any, all, or none of the configured views. The hardware configuration detailed in table 1 was successfully tested as part of the work.

Table 1 Example VirtualwindoW Configuration

View Number	Included Devices
1	Logitech Head Tracker, Pan and Tilt #1 with Stereo Camera Pair #1
2	Logitech Head Tracker, Pan and Tilt #2 (inverted) with Stereo Camera Pair #2
3	PC Keyboard Motion Controller, Pan and Tilt #3 with single Sony Zoom Camera
4	Logitech Head Tracker, Multiplexed Pan and Tilt #4 and IT Zoom run over single serial Radio Link, video signal over separate radio link
5	Logitech Head Tracker, Small Robot controlled via radio modem, single fixed camera video signal over separate radio link

To be able to support cases where several identical pieces of hardware are connected (i.e. the four PTU sets above) the software was also modified so that one piece runs them all. Instead of using separate code to contain the configuration of particular devices, now a standard system data structure loaded from the plain text configuration file is used. It is also possible to provide parameters to these modules to change their behavior. A good example of this is being able to use a PTU in the physically inverted position. The only difference between an inverted and non-inverted PTU is the requirement to move opposite from each other for the same input.

3.3 Generalized Zone Concept

When the zoom cameras were added to the VirtualwindoW system, there was a need to express this in a way consistent with the existing head-tracking method. To do this, an intuitive lean forward/backwards motion was devised to initiate increasing and decreasing zoom amounts respectively. One of the zoom cameras also required the ability to manually focus the camera. This was implemented through the motion of the operator’s head tilting from side to side. The complete set of supported head motions is shown in Figure 3. These movements correspond to traditional x, y, z, yaw, pitch, and roll motions except that the z motion has not been used.

Supporting all of these inputs required a way to easily represent them internally. To do this, INEEL researchers developed the “generalized zone” concept. The generalized zone takes each input command (zoom, focus, tilt, pan, and slide) and divides it into graduated sections like before, with each one now represented by an integer code. The sign of this code is used to indicate direction and increasing magnitude corresponds to faster motion. Testing this approach with zooming devices revealed a problem. When viewing at high magnification, even the slowest pan or tilt speed would completely lose track of the area of interest. To allow the motion speed to be graduated, it was necessary to add an additional code for magnification factor, which is generated in a zoom device when present. Increasing values of this code indicate that the current magnification factor measured at the lens is higher.

All of these codes are packaged within a standard structure that is then passed through all devices in use for the current view. It is then up to each device module how to act upon that information. If a particular type of head motion does not make sense for a piece of hardware, it is ignored. For example, a PTU only uses pan, tilt and magnification, an auto-focusing zoom camera only uses the zoom code, and a linear slider bar only uses the slide code. It is even possible to use the codes for intuitively similar functions. A later section will discuss the remote control of a mobile robot using generalized zones.

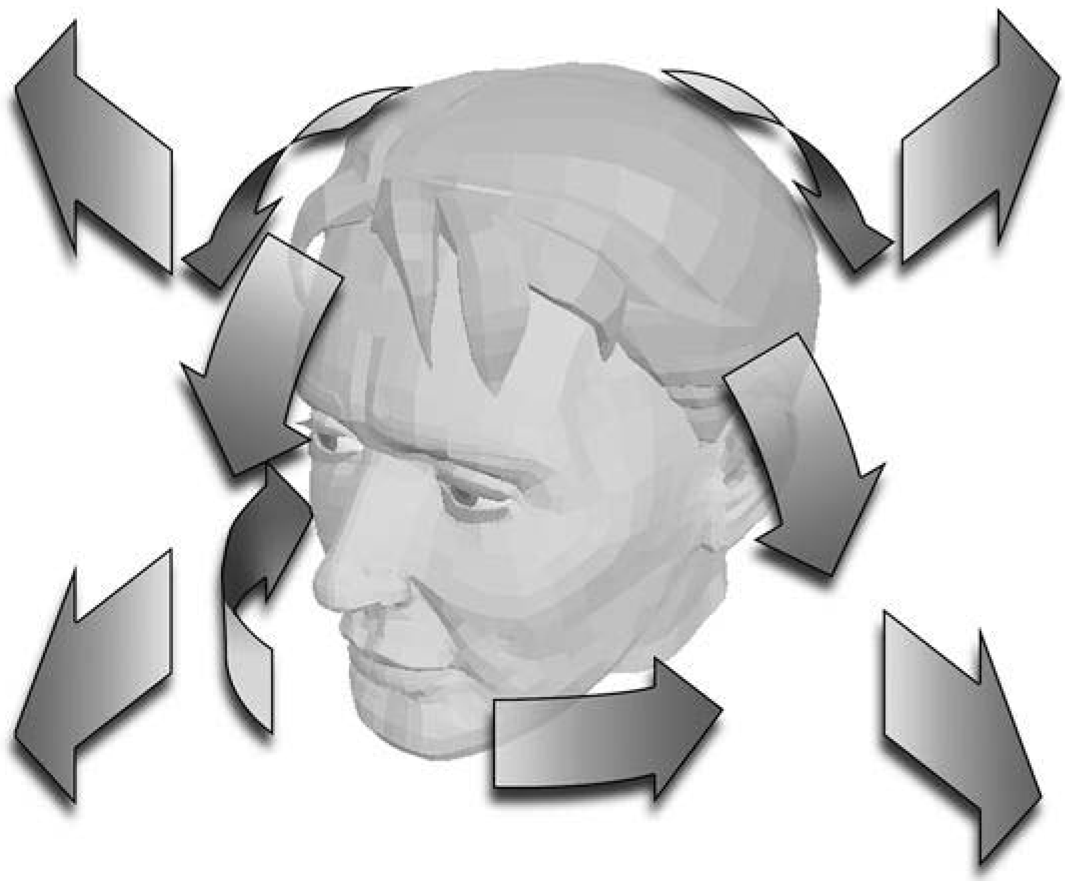

Figure 3 VirtualwindoW Head Motions

The computations required to generate the codes are performed within the software module for the main input device of the current view. Earlier software allowed only head-tracker input, but using the zone codes to communicate with other system modules makes it much easier to add other devices. It took only a small amount of time to add a keyboard control module that can be used in place of the head tracker. Other possible input devices include joysticks or three-dimensional mice. The addition of new moving hardware devices is also simplified because their software modules need only react to the generalized zone codes they receive. The computation of those codes is completely independent.

3.4 New Diagnostic Tools

The new modular architecture made it easier to add diagnostic capabilities that previously required separate software. Most current hardware has been connected to the VirtualwindoW system using RS-232 serial ports. Now that there are eight ports available for use, it was found that standard off-the-shelf terminal software would generally not support communications with all of them. A text terminal was often found to be helpful in the process of developing for new serial-based devices or troubleshooting existing ones. Since the system already has the capability to configure and use these ports, a software module was created to provide a built-in text terminal. When this diagnostic terminal module is included in the current selected view the lower part of the VirtualwindoW diagnostic display will show any characters received from that device and characters typed on the keyboard are sent straight out the serial port. This is useful for field troubleshooting and removes the need to use external serial terminal software in new module development.

With the many different generalized zone codes implemented in the system, a way of visually representing them was needed. To this end, a screen display device was created that is able to represent the zoom, focus, tilt, pan, and slide in a form that corresponds roughly to head position. In this case, all that this small display module does is show an indicator on the screen based on the zone values. Another diagnostic device was implemented that shows the current values of all zone structure codes (including magnification) as they change for more in-depth diagnosis.

3.5 New COTS Hardware

Once all of these changes were made to the computer system and software, it was a straightforward matter to add support for additional hardware devices. This was very useful because several different avenues were explored to provide the stereo zoom capability that was the basis for the latest VirtualwindoW research. This investigation required purchasing several commercial components for integration into the system.

Figure 4 Sony Block on PTU

The first camera investigated was the Sony model EVI-370DG[3]. One of these cameras is shown in figure 4. This self-contained camera block was originally designed for use in consumer video camera/recorders (camcorders). Several features made it a strong candidate for use in stereo zooming: Gen-Lock video capability, RS-232 serial interface control, automatic focus, and the ability to synchronize its zoom with another camera. For our investigation, two of these devices were used in place of a dual camera set we had already been using successfully. The camera blocks were configured so that they would zoom together and provide Gen-Locked video signals needed by the Crystal Eyes stereo vision system.

The video quality and zoom functionality of this camera was adequate for this application. Unfortunately it was not a viable solution due to physical geometry considerations. The Crystal Eyes system requires the two cameras providing input be separated by a particular distance and angle. When the magnification factor was changed while using the Sony cameras, small adjustments in separation were required to produce an acceptable stereo view. It was not feasible within the scope of this project to create a method to properly vary the camera separation with the magnification factor so this approach was abandoned. However, since the development had already been done to control the camera, a single camera unit was incorporated into the system and found useful for non-stereo viewing.

The next approach attempted to solve the stereo zoom problem using a novel single camera approach. International Telepresence (IT) has developed a method to provide the two separate video signals from a single video camera using some added proprietary optical hardware and Digital Signal Processing (DSP). This device had already seen successful use in medical applications such as endoscopic surgery[4] that require magnification and stereo vision together.

One of these IT systems was custom configured, purchased, and integrated with a separate zoom lens produced by the Rainbow Corporation. The Rainbow H6 x 8M-II provided a maximum magnification factor of six, manual focus control, and the ability to measure the current zoom amount. A small microcontroller board was added to control the zoom lens based on RS-232 serial commands from the VirtualwindoW system. This board was also responsible for measuring the current magnification factor and returning it to the system for use in computing the generalized magnification zone code.

Figure 5 IT Zoom Block on PTU

The camera unit shown in figure 5 is the Hitachi KP-D50. It was chosen because it has integrated DSP capabilities that can enhance the brightness of the video signal it perceives. This was necessary because one cost of using the International Telepresence stereo system is a reduced amount of light making it through all of the optical components to the camera. Indoor pictures were often dark and grainy before using this camera unit. Another cost was the large weight of the total assembly. The camera, IT stereo device, and zoom lens were too heavy for the PTU we had been using. Directed Perception, manufacturer of the PTU in use on the system, also produces a physically identical device that trades increased weight capacity (torque) for movement speed. This unit also uses the same serial command language, which meant that no software changes were needed (although it would have been straightforward to add a module had it been necessary).

In this configuration, the International Telepresence-based system was found to be very useable for applications requiring depth perception and magnification. Use of the head-tracking system to control magnification was found to be both intuitive and accurate enough to be useable. One drawback of this particular arrangement is the requirement for the operator to control the focus, but that was also resolved through use of the new focus head motion as described above in the section on generalized zones.

4. UNMANNED VEHICLE APPLICATIONS

The INEEL Robotics Group has been conducting unmanned vehicle research in cooperation with Utah State University for several years. After the VirtualwindoW stereo zoom investigation was completed, additional work specifically applicable to this area was performed. Wireless control, vehicle operation, and on-vehicle stereo viewing were all areas of examination.

4.1 Wireless Device Control

For optimum mobility, remote vehicles operate without a cable or tether to the control station. VirtualwindoW has until now operated solely with cable connections to the various devices connected to it. As discussed, the initial work in 1995 was conducted on a tethered, tracked vehicle. In order to be successfully applied to the mobile vehicle arena, the VirtualwindoW must also be able to control its devices through a wireless interface. Since most of the devices were already controlled using RS-232 serial interfaces, the first step was substituting radio frequency modems in place of cables. DGR-115 wireless modems built by Freewave were used in place of the communications cables for the PTU and Rainbow zoom lens setup detailed above. This functioned well—but required two pairs of modems.

In an effort to reduce this to a single link, modifications were made to both the Rainbow zoom microcontroller and to the main VirtualwindoW software. Instead of having its own radio link, the PTU was connected to a spare serial port on the microcontroller board that runs the zoom lens. A new combined device was added to the VirtualwindoW system that sends the communications streams needed for both the PTU and the Rainbow Zoom over a single serial port. These separate streams are encapsulated within packets and sent from the control station. This method was also successful, and is appropriate when there is a computing unit on the remote side as there was in this example.

4.2 Wireless Video

To make VirtualwindoW completely wireless, video signals must also be transmitted to the control station. Microtek Minilink 2.4GHz short-range video transmission sets were used to test the usability of remote video. The number of links required varies with both the video type and whether stereo or non-stereo viewing is required. For this research, INEEL investigated the use of clearer, crisper S-Video as well as standard NTSC video signal protocols. An NTSC video stream requires a single Microtek wireless link to transmit. INEEL discovered a novel way to use two Microtek links to transmit a single S-Video signal successfully. Table 2 describes the number of wireless video links required for various camera configurations.

Table 2 Video Link Configurations

Description of Video Configuration	Number of Links Required
Non-Stereo NTSC Video	1
Dual-Camera Stereo over NTSC Video	2
Non-Stereo S-Video	2
Dual Camera Stereo over S-Video	4
IT Stereo over NTSC Video	1
IT Stereo over S-Video	2

While the wireless transmitters functioned properly, it was discovered that the video transmission quality varied greatly as the remote camera was moved around, even at short range. This was due to the interference from metal objects and other electromagnetic sources. The low transmission power was not sufficient to maintain a consistent video sync under these circumstances, which manifested itself in the VirtualwindoW screen picture dropping out. The large, high quality video monitor used to view system video would take several seconds to re-synchronize every time this happened. These pauses were long enough to disrupt the operator's control of the device. This could be improved by using a better short-haul video transmitter, perhaps one with more power, but this is often not an option with smaller vehicles or with ones that use battery power.

4.3 Vehicle Control

The intuitive head motions used in the VirtualwindoW system can be extended to control of a vehicle itself. A brief test was conducted using a platform called the Advanced Robot Chassis (ARC) III. This vehicle measures 24 inches long, 18 inches wide, and 18 inches high (see figure 6). It has six independently controlled wheels that allow for very flexible motion. It is controlled using a digital serial protocol over a radio link and has a single fixed video camera and transmitter mounted to view the area in front of the vehicle. The development of the ARC III is fully described in a paper presented to the SPIE[5].

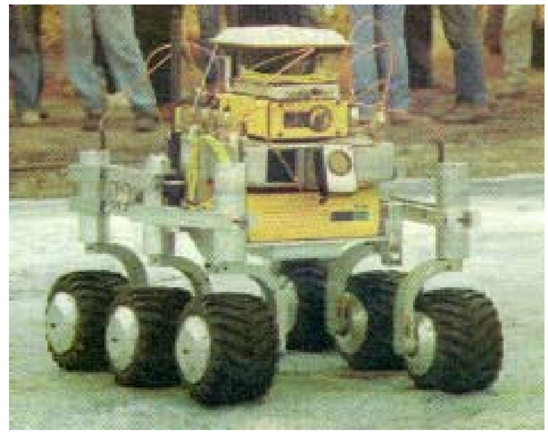

Figure 6 ARC III Vehicle

The communications protocol used to control this robot was built into a new VirtualwindoW module and a compatible radio transmitter was connected. Head-tracker motions were picked that would provide intuitive vehicle control. Leaning forward and backward, which would normally change the magnification factor, instead moved the ARC forward and backward. The side-to-side slider motion was adapted to translate the vehicle sideways. The pan left and right motion was adapted to turning the vehicle.

The short test of this concept was successful, with some reservations. The head motions required some getting used to, and some tuning of the relative amount of head motion required was needed for varying speeds. There was also a time lag between head motion and the response of the vehicle. The video dropout problem discussed above proved to be the largest problem. Frequently the video synchronization would be lost, causing the video monitor to momentarily blank. Since the video is the operator's motion control feedback, smooth control was sometimes disrupted. This was particularly evident during turn maneuvers. Although the low power video transmitter caused some operational difficulties, the concept of controlling a mobile device was proven and with some minor enhancements to the system smooth and accurate control can be obtained.

Figure 7 RATV Vehicle

It should be noted that this work can be immediately applied to larger vehicles. INEEL is currently working with another vehicle called the Remote All Terrain Vehicle (RATV) which is a 4-passenger ATV commercially available from Triton. The RATV has been converted to remote operation for testing as an agricultural and environmental management automation platform[6] (see figure 7). This vehicle uses the same remote control protocol as the ARC III, but has much longer range and endurance due to it internal combustion propulsion.

Work is currently underway in cooperation with Utah State University to test the RATV and ARC III in cooperative operation scenarios. One RATV has been modified to deploy and retrieve an ARC III for field deployment as part of this effort.

5. CONCLUSION

From its inception in 1995 as a single purpose device, the INEEL VirtualwindoW has grown into a modular, flexible, system for deploying multiple video perception technologies. This single platform can control and display multiple mono and stereo video inputs in a hands-free manner. Improvements such as the generalized zone concept, the human-readable configuration file, and new diagnostic tools have been added. These changes have made it a simple matter to add support for new commercial hardware devices to the system software, and several were during the course of this work. Some of those components were added particularly to support investigation into the suitability of VirtualwindoW technology to remote vehicle applications. Examples of this are the use of wireless device control and adapting the generalized zone concept to directly controlling the motion of a mobile robotic vehicle. The VirtualwindoW has proven to be useful in this area, and merits refinement and continuation of this work.

ACKNOWLEDGEMENTS

INEEL's Academic and Professional Affairs Department provided research funding for this project through manager Debonny Shoaf with Laboratory Directed Research and Development staff Diana Skoy and Steven Piet. The ARC III and RATV autonomous vehicles were built at Utah State University in collaboration with the INEEL. Mark Paulsen provided useful graphic arts services for this paper. The assistance of the International Telepresence Company in the adaptation of their single camera stereo vision system is greatly appreciated.

REFERENCES

1. McKay, M.D. and Anderson, M.O., "Telepresence For Mobile Robots in Nuclear Environments", presented at the International Society for Optical Engineering Symposium, Boston, MA, November, 1996.
2. Anderson, M.O., McKay, M.D., and Willis, W.D., "The VirtualwindoW for Nuclear Applications", presented at the American Nuclear Society 7th Topical Meeting on Robotics and Remote Systems, Augusta, Georgia, April, 1997.
3. Sony Corporation, "EVI-370DG Color Camera Block User Manual"
4. Fountoulakis, A., Lomax, A., and Chadwick, S.J.D., "Initial Experience of 3-D Video Endoscopy in General Surgery", Minimal Access Surgery – Surgical Technology International V, pp.115-119.
5. Poulson, E.A., Jacobs, J.S., Gunderson, R.W., and Abbot, B., " Design of a Robotic Vehicle with Self-Contained Intelligent Wheels", Proceedings of the SPIE Conference on Robotic and Semi-Robotic Ground Vehicle Technology, Orlando, Florida, April 1998, pp. 68-73.
6. Jacob, J.S., Gunderson, R.W., Fullmer, R.R., "Conversion and Control of an All-Terrain Vehicle for Use as an Autonomous Mobile Robot", Proceedings of the SPIE Conference on Robotic and Semi-Robotic Ground Vehicle Technology, Orlando, Florida, April 1998, pp.83-92.